

# Erdős–Szekeres-type statements: Ramsey function and decidability in dimension 1\*

BORIS BUKH

Centre for Mathematical Sciences  
Wilberforce Rd, University of Cambridge  
Cambridge CB3 0WB, Great Britain and

Churchill College, Storey’s Way  
Cambridge CB3 0DS, Great Britain

JIŘÍ MATOUŠEK

Department of Applied Mathematics  
Charles University, Malostranské nám. 25  
118 00 Praha 1, Czech Republic, and

Institute of Theoretical Computer Science  
ETH Zurich, 8092 Zurich, Switzerland

## Abstract

A classical and widely used lemma of Erdős and Szekeres asserts that for every  $n$  there exists  $N$  such that every  $n$ -term sequence  $\underline{a}$  of real numbers contains an  $n$ -term increasing subsequence or an  $n$ -term nondecreasing subsequence; quantitatively, the smallest  $N$  with this property equals  $(n - 1)^2 + 1$ . In the setting of the present paper, we express this lemma by saying that the set of predicates  $\Phi = \{x_1 < x_2, x_1 \geq x_2\}$  is Erdős–Szekeres with Ramsey function  $\text{ES}_\Phi(n) = (n - 1)^2 + 1$ .

In general, we consider an arbitrary finite set  $\Phi = \{\Phi_1, \dots, \Phi_m\}$  of *semialgebraic predicates*, meaning that each  $\Phi_j = \Phi_j(x_1, \dots, x_k)$  is a Boolean combination of polynomial equations and inequalities in some number  $k$  of real variables. We define  $\Phi$  to be *Erdős–Szekeres* if for every  $n$  there exists  $N$  such that each  $N$ -term sequence  $\underline{a}$  of real numbers has an  $n$ -term subsequence  $\underline{b}$  such that at least one of the  $\Phi_j$  holds everywhere on  $\underline{b}$ , which means that  $\Phi_j(b_{i_1}, \dots, b_{i_k})$  holds for every choice of indices  $i_1, i_2, \dots, i_k$ ,  $1 \leq i_1 < i_2 < \dots < i_k \leq n$ . We write  $\text{ES}_\Phi(n)$  for the smallest  $N$  with the above property.

We prove two main results. First, the Ramsey functions in this setting are at most doubly exponential (and sometimes they are indeed doubly exponential): for every  $\Phi$  that is Erdős–Szekeres, there is a constant  $C$  such that  $\text{ES}_\Phi(n) \leq 2^{2^{Cn}}$ . Second, there is an algorithm that, given  $\Phi$ , decides whether it is Erdős–Szekeres; thus, one-dimensional Erdős–Szekeres-style theorems can in principle be proved automatically.

We regard these results as a starting point in investigating analogous questions for  $d$ -dimensional predicates, where instead of sequences of real numbers, we consider sequences of points in  $\mathbb{R}^d$  (and semialgebraic predicates in their coordinates). This setting includes many results and problems in geometric Ramsey theory, and it appears considerably more involved. Here we prove a decidability result for *algebraic* predicates in  $\mathbb{R}^d$  (i.e., conjunctions of polynomial equations), as well as for a multipartite version of the problem with arbitrary semialgebraic predicates in  $\mathbb{R}^d$ .

## 1 Introduction

**Motivation and background.** Ramsey-type theorems claim that, generally speaking, any sufficiently large structure of a given kind contains a “very regular” substructure of a prescribed size. The following two gems from a 1935 paper of Erdős and Szekeres [7] belong among the earliest, best known, and most useful instances.

---

\*Both authors were supported by the ERC Advanced Grant No. 267165.

**Theorem 1.1** (On monotone subsequences). *For every  $n \geq 2$ , every sequence  $(a_1, a_2, \dots, a_N)$  of real numbers, with  $N \geq (n-1)^2 + 1$ , contains a monotone subsequence of length  $n$ ; more precisely, there are indices  $i_1 < i_2 < \dots < i_n$  such that either  $a_{i_1} < \dots < a_{i_n}$  or  $a_{i_1} \geq \dots \geq a_{i_n}$ .*

See, for example, Steele [17] for a collection of six nice proofs and many applications.

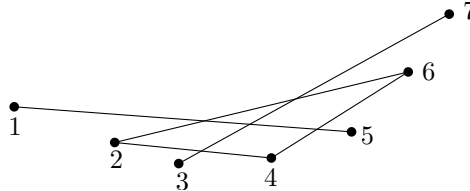
**Theorem 1.2** (On subsets in convex position). *For every  $n \geq 3$ , among every  $N = \binom{2n-4}{n-2} + 1 \approx 4^n/\sqrt{n}$  points in the plane, no three collinear, one can always select  $n$  points in convex position (i.e., forming the vertex sets of a convex  $n$ -gon).*

See, e.g., [12, 10] for proofs and surveys of developments around this result.

Many geometric Ramsey-type questions have been investigated in the literature; here we mention just three examples that directly motivated our research.

The *colored Tverberg theorem*, conjectured by Bárány, Füredi, and Lovász [1] and proved by Vrećica and Živaljević [20] asserts that *for every  $d$  and  $r$  there exists  $t$  such that if  $A_1, \dots, A_{d+1}$  are  $t$ -point sets in  $\mathbb{R}^d$  (we imagine that each  $A_i$  has its own color; e.g., the points of  $A_1$  are red, those of  $A_2$  blue, etc.), there are  $r$ -point subsets  $B_i \subseteq A_i$ ,  $i = 1, 2, \dots, d+1$  and a point  $\mathbf{x} \in \mathbb{R}^d$  that lies in the convex hull of  $\{\mathbf{b}_1, \dots, \mathbf{b}_{d+1}\}$  for every choice of  $\mathbf{b}_1 \in B_1, \dots, \mathbf{b}_{d+1} \in B_{d+1}$ .* For us, this result is remarkable because all known proofs use topological methods—there is no known proof by “ordinary” geometric and/or combinatorial arguments (although some special cases do have elementary proofs).

The following result was needed as a lemma in the paper [3] by Loh, Nivasch, and the first author: *every sequence  $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N)$  of points in  $\mathbb{R}^2$  with increasing  $x$ -coordinates has a subsequence  $(\mathbf{b}_1, \dots, \mathbf{b}_n)$  in which every 7-tuple  $\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_7}$ ,  $i_1 < \dots < i_7$ , is as in the next picture, i.e., the triangle  $\mathbf{b}_{i_2}\mathbf{b}_{i_4}\mathbf{b}_{i_6}$  contains the intersection of the segments  $\mathbf{b}_{i_1}\mathbf{b}_{i_5}$  and  $\mathbf{b}_{i_3}\mathbf{b}_{i_7}$ —provided that  $N$  is sufficiently large in terms of  $n$ .*



While the proof in [3] is simple, the situation with an appropriate  $d$ -dimensional generalization is discouraging: there are increasingly complicated proofs up to dimension 4, while the 5-dimensional case already seems out of reach with the present methods.

Yet another example we want to mention is a Ramsey-type result of Eliáš and the second author [6]: *every  $N$ -point sequence  $(\mathbf{a}_1, \dots, \mathbf{a}_N)$  of points in the plane with increasing  $x$ -coordinates has an  $n$ -term subsequence that is  $k$ th order monotone, meaning that either every  $(k+1)$ -tuple of points lies on the graph of a smooth function with nonnegative  $k$ th derivative, or every  $(k+1)$ -tuple lies on the graph of a smooth function with nonpositive  $k$ th derivative.* (This is a common generalization of Theorems 1.1 and 1.2.) Here  $N$  depends on  $n$  and  $k$ ; the existence of *some*  $N$  follows immediately from Ramsey’s theorem for  $(k+1)$ -tuples, but an interesting question here is the behavior of the Ramsey function—how big is the smallest  $N = N_k(n)$  that works. We have  $N_1(n) \approx n^2$  and  $N_2(n) \approx 4^n$  according to Erdős and Szekeres, and in [6] it was proved that  $N_3(n)$  is doubly exponential, i.e.,  $2^{2^{c_1 n}} \leq N_3(n) \leq 2^{2^{c_2 n}}$ . The

order of magnitude of  $N_4(n)$  is unknown, and so are the Ramsey functions for numerous other geometric Ramsey-type results.

**Erdős–Szekeres predicates.** The above examples and some others motivate (at least) three general questions formulated below. Before stating them, we introduce some notions and notation.

Let  $k$  be an integer, which we think of as small and fixed, and let  $\Phi = \Phi(\mathbf{x}_1, \dots, \mathbf{x}_k)$  be a  $d$ -dimensional  $k$ -ary predicate, by which we mean a mapping  $(\mathbb{R}^d)^k \rightarrow \{\text{False}, \text{True}\}$ . We say that  $\Phi$  *holds everywhere* on a sequence  $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  of points in  $\mathbb{R}^d$  if  $\Phi(\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_k})$  holds for every increasing  $k$ -tuple  $i_1, \dots, i_k$  of indices,  $1 \leq i_1 < i_2 < \dots < i_k \leq n$ .

**Definition 1.3.** Let  $\Phi$  be a set of  $d$ -dimensional predicates. We say that  $\Phi$  is Erdős–Szekeres if for every  $n$  there exists  $N$  such that every sequence  $\mathbf{a}$  of length  $N$  has subsequence  $\mathbf{b}$  of length  $n$  such that there is a predicate  $\Phi \in \Phi$  that holds everywhere on  $\mathbf{b}$ . Let  $\text{ES}_\Phi(n)$  denote corresponding Ramsey function, i.e. the smallest  $N$  with the property above.

For example, Theorem 1.1 on monotone subsequences can be re-stated as follows in this language: the set  $\Phi = \{x_1 < x_2, x_1 \geq x_2\}$  of 1-dimensional predicates is Erdős–Szekeres, with  $\text{ES}_\Phi(n) = (n-1)^2 - 1$  (note that we do not write points in  $\mathbb{R}^1$ , i.e., real numbers, in boldface—unlike points in  $\mathbb{R}^d$ ).

For Theorem 1.2 on subsets in convex position, we can let  $\Phi$  consist of a single 4-ary predicate  $\Phi_{\text{conv}}(\mathbf{x}_1, \dots, \mathbf{x}_4)$  expressing that the 4-tuple  $\mathbf{x}_1, \dots, \mathbf{x}_4$  is in convex position (with appropriate handling of collinearities). Indeed, a set is in convex position if each of its 4-tuples is. Alternatively, we can set  $\Phi' := \{\Phi_{\text{pos}}, \Phi_{\text{neg}}\}$ , where  $\Phi_{\text{pos}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$  expresses that the ordered triple  $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$  has a positive orientation, while  $\Phi_{\text{neg}}$  expresses negative orientation.

Let us remark that the notion of Erdős–Szekeres set of predicates is not general enough to express the colored Tverberg theorem mentioned above, for example; however, it could easily be extended to a colored setting, if that proved useful.<sup>1</sup>

Another thing worth noting here is that given a finite set  $\Phi$  of predicates, it is possible to produce a single predicate  $\Phi$  such that  $\Phi$  is Erdős–Szekeres iff  $\{\Phi\}$  is. Moreover, the functions  $\text{ES}_\Phi$  and  $\text{ES}_{\{\Phi\}}$  can be related; see Section 8 for a discussion. However, the passage from  $\Phi$  to  $\Phi$  increases the number of variables and produces a rather cumbersome predicate  $\Phi$ , while treating a set of predicates in our development is not much more complicated than treating a single one, so we prefer dealing with sets of predicates.

The predicates in Theorems 1.1 and 1.2, as well as in the other examples above and elsewhere in geometric Ramsey theory, can be represented as semialgebraic predicates, where a  $d$ -dimensional  $k$ -ary semialgebraic predicate  $\Phi = \Phi(\mathbf{x}_1, \dots, \mathbf{x}_k)$  is a Boolean combination of polynomial equations and inequalities in the coordinates of the  $k$  points  $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d$  (the polynomials are assumed to have rational coefficients). In this paper we consider only semialgebraic predicates.

We consider the following general questions quite natural and very interesting.

<sup>1</sup>Perhaps more seriously, there are results and problems in geometric Ramsey theory that do not fit this kind of framework at all. For example, the famous recent *6-hole theorem* of Gerken [8] and Nicolás [13] asserts that every sufficiently large point set  $P \subset \mathbb{R}^2$  in general position contains an empty hexagon, i.e., six points in convex position whose convex hull contains no other point of  $P$ . Here the problem is that this is not a “pure” Ramsey-type result, since, although we pass to a subset, the original points still play a role as “obstacles”. We suspect general problems of this kind to be substantially harder than those considered in this paper.

1. (Ramsey function) *What is the largest possible order of magnitude of  $\text{ES}_\Phi$ , where  $\Phi$  is a finite set of  $d$ -dimensional  $k$ -ary semialgebraic predicates?*

A simple upper bound follows from Ramsey's theorem. For  $\Phi = \{\Phi_1, \dots, \Phi_m\}$  and a sequence  $\mathbf{a}$  of length  $N$ , we color a  $k$ -tuple  $I \subseteq \{1, 2, \dots, N\}$  with color  $i \in \{1, 2, \dots, m\}$  if  $\Phi_i$  holds for the  $k$ -tuple in  $\mathbf{a}$  indexed by  $I$  (if there are several possibilities, we pick one arbitrarily). The color  $m + 1$  is used for the  $k$ -tuples on which no  $\Phi_i$  holds. By Ramsey's theorem, if  $N \geq R_k(n; m + 1)$  (the Ramsey number for  $k$ -tuples with  $m + 1$  colors),  $\mathbf{a}$  contains a subsequence of length  $n$  in which all  $k$ -tuples have the same color. If  $\Phi$  is Erdős–Szekeres, then for  $n$  sufficiently large, color  $m + 1$  is impossible. Thus, we get  $\text{ES}_\Phi(n) \leq R_k(n; m + 1)$  for all sufficiently large  $n$ , and  $R_k(n; m + 1)$  is bounded above by a tower function of height  $k$ —see, e.g., [9].

2. (Decidability) *Is there an algorithm that, given a finite set  $\Phi$  of  $d$ -dimensional semialgebraic predicates, decides whether  $\Phi$  is Erdős–Szekeres?*

By a celebrated result of Tarski [18], the first-order theory of the reals is decidable. That is, there is an algorithm deciding the validity of formulas of the form

$$(Q_1 x_1) \dots (Q_k x_k) \Phi(x_1, \dots, x_k),$$

where  $Q_1, \dots, Q_k \in \{\forall, \exists\}$  are quantifiers for the real variables  $x_1, \dots, x_k$  and  $\Phi$  is a semialgebraic predicate in our sense. This is one of the most useful decidability results, and there is an extensive area studying efficient algorithms for this problem and various special cases (see, e.g., [2]). This may give some hope for decidability of the Erdős–Szekeres property, and quantifier-elimination methods for the first-order theory of the reals play a significant role in the present paper.

3. (Homogeneous, or indiscernible, sequences) *What can be said about very long, or infinite, point sequences that are homogeneous w.r.t. interesting classes of semialgebraic predicates?*

To explain this question, let us first consider all 2-dimensional semialgebraic predicates  $\Phi$  that depend only on the orientation of triples of points; for example, “ $\mathbf{x}_1$  lies in the convex hull of  $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_3$ ”. Then every point sequence  $C$  in convex position (i.e., the points are in convex position and numbered along the circumference of the convex hull either clockwise or counterclockwise) is *homogeneous* for this class of predicates, meaning that each  $k$ -ary such predicate  $\Phi$  either holds on all  $k$ -tuples of  $C$  or holds on none of them. Moreover, as far as these predicates are concerned, all sequences in convex position look the same (except for two possible orientations, that is).

However, some natural Ramsey-type questions, such as the colored Tverberg problem or the question with triangles containing segment intersections illustrated by the picture above, need “higher-order” geometric predicates, such as whether the intersection of two lines, each determined a pair of the considered points, lies above or below the line spanned by another pair of points. (Semialgebraic predicates of course include this kind of predicates and much more.) Here it is far from obvious what homogeneous sequences for such predicates might look like. Let us remark that one kind of a homogeneous set for semialgebraic predicates of bounded complexity, called the *stretched diagonal*, was used as an interesting example in the recent papers [4, 5], together with a related

construction of a *stretched grid*—this is yet another motivation of our investigations here.

**Answers for dimension 1.** Luckily, after investigating this kind of problems for some time, we found the (unpublished) thesis of Rosenthal [16] which, in the language of logic and model theory, provided an answer to the third question (homogeneous sets) for  $d = 1$ . With the help of Rosenthal’s results and methods, together with other techniques, we then succeeded in answering the first two questions for  $d = 1$ .

We were surprised by the answer to the first question: it turned out that in dimension 1, the Ramsey functions can be at most doubly exponential, independent of the arity of the predicates.

**Theorem 1.4.** *For every finite Erdős–Szekeres set  $\Phi$  of 1-dimensional predicates there exists a number  $C = C(\Phi)$  such that*

$$\text{ES}_{\Phi}(n) \leq 2^{2^{Cn}}.$$

The proof is presented in Sections 2–4.

In Section 7 we will show that some predicates indeed require a doubly exponential bound.

**Proposition 1.5.** *There is a finite Erdős–Szekeres set  $\Phi$  of predicates with  $\text{ES}_{\Phi}(n) \geq 2^{2^{cn}}$ , with a suitable constant  $c > 0$ .*

The proof of Theorem 1.4 also immediately provides a quantitative result for homogeneous subsequences for every semialgebraic predicate, Erdős–Szekeres or not.

**Proposition 1.6.** *For every finite set  $\Phi$  of 1-dimensional semialgebraic predicates there is a constant  $C = C(\Phi)$  such that for every integer  $n$ , every sequence  $\underline{a}$  of length  $2^{2^{Cn}}$  contains an  $n$ -term subsequence  $\underline{b}$  such that for each predicate  $\Phi \in \Phi$  holds either everywhere on  $\underline{b}$  or nowhere on  $\underline{b}$ .*

By adding some more ingredients to the proof of Theorem 1.4 in Sections 5 and 6, we will also prove that 1-dimensional Erdős–Szekeres predicates can be recognized algorithmically.

**Theorem 1.7.** *There is an algorithm that, given a finite set  $\Phi$  of 1-dimensional semialgebraic predicates, decides whether it is Erdős–Szekeres.*

**Higher dimensions: algebraic predicates and the multipartite case.** For  $d \geq 2$ , all of the three questions above appear much harder than for  $d = 1$ , and at the time of writing this paper we have just some preliminary results for  $d = 2$ .

Here we provide decidability results with  $d$  arbitrary for a restricted class of predicates, as well as for general semialgebraic predicates but with a different Ramsey-type question.

Let us define an *algebraic predicate*, a special kind of a semialgebraic predicate, as a  $d$ -dimensional predicate  $\Phi(\mathbf{x}_1, \dots, \mathbf{x}_k)$  expressible as a conjunction of polynomial equations in the coordinates of  $\mathbf{x}_1, \dots, \mathbf{x}_k$  (with rational coefficients). We note that while a semialgebraic predicate defines a semialgebraic set (in  $(\mathbb{R}^d)^k$ ), an algebraic predicate defines an algebraic variety.

For algebraic predicates the question of being Erdős–Szekeres does not make much sense, since every nontrivial algebraic predicate fails on a generic  $k$ -tuple of points. However, the question of whether there exist arbitrarily long point sequences, or even infinite ones, on which a given algebraic predicate holds everywhere, is meaningful. The following theorem shows that it is decidable.

**Theorem 1.8** (“Effective compactness” for algebraic predicates). *For every  $d$ ,  $D$ , and  $k$  there exists  $N$ , for which an explicit bound can be given, such that for every  $d$ -dimensional  $k$ -ary algebraic predicate  $\Phi$ , in which all the polynomials have degree at most  $D$ , the following two conditions are equivalent:*

- (i) *There exists a sequence  $\underline{a}$  of  $N$  points in  $\mathbb{R}^d$  on which  $\Phi$  holds everywhere.*
- (ii) *There exists an infinite sequence  $\underline{a}$  of points in  $\mathbb{R}^d$  on which  $\Phi$  holds everywhere.*

Since condition (i) can be tested using a decision algorithm for the first-order theory of the reals, we obtain a decision algorithm for testing the existence of infinite sequences, or equivalently, of arbitrarily long sequences, on which a given algebraic predicate holds everywhere.

Let us remark that the statement of the theorem also holds with the real numbers replaced by the complex numbers.

We note that a statement analogous to Theorem 1.8 fails badly for semialgebraic predicates. Indeed, the 1-dimensional binary predicate  $\Phi(x_1, x_2) := (x_1 > 0) \wedge (x_2 \geq x_1 + 1) \wedge (x_2 \leq 2x_1)$  holds everywhere on arbitrarily long finite sequences, such as  $(n, n+1, \dots, 2n)$ , but on no infinite sequence. Similarly,  $(x_1 > 0) \wedge (x_2 \geq x_1 + 1) \wedge (x_2 \leq A)$ , where  $A$  is a constant, admits  $A$ -term sequences but not longer.

Next, we let  $\Phi(\mathbf{x}_1, \dots, \mathbf{x}_k)$  be a  $d$ -dimensional semialgebraic predicate, but we ask a different question. Let  $A_1, \dots, A_k$  be point sets in  $\mathbb{R}^d$ . We say that  $\Phi$  holds everywhere on  $A_1 \times \dots \times A_k$  if  $\Phi(\mathbf{a}_1, \dots, \mathbf{a}_k)$  holds for every choice of points  $\mathbf{a}_1 \in A_1, \dots, \mathbf{a}_k \in A_k$ . We have the following analog of Theorem 1.8.

**Theorem 1.9** (“Effective compactness” for the multipartite setting). *For every  $d$ ,  $D$ , and  $k$  there exists  $N$ , for which an explicit bound can be given, such that for every  $d$ -dimensional  $k$ -ary semialgebraic predicate  $\Phi$ , in which all the polynomials have degree at most  $D$ , the following two conditions are equivalent:*

- (i) *There exist  $N$ -point sets  $A_1, \dots, A_k$  in  $\mathbb{R}^d$  such that  $\Phi$  holds everywhere on  $A_1 \times \dots \times A_k$ .*
- (ii) *There exist infinite sets  $X_1, \dots, X_k$  in  $\mathbb{R}^d$  such that  $\Phi$  holds everywhere on  $X_1 \times \dots \times X_k$ .*

Theorems 1.8 and 1.9 are proved in Section 9; the proofs are similar and more or less independent of the rest of the paper.

## 2 Ramseying for fast-growing sequences

Here we work in the 1-dimensional setting, i.e., with sequences of real numbers, and we begin with preparations for the proofs of Theorems 1.4 and 1.7.

**$R$ -growing sequences.** The first idea in our approach is that if  $\Phi$  is a 1-dimensional semialgebraic predicate and  $\underline{a}$  is a sequence that grows sufficiently fast (where the speed of growth is quantified with respect to  $\Phi$ ), then either  $\Phi$  holds everywhere on  $\underline{a}$ , or  $\Phi$  holds nowhere on  $\underline{a}$  (by which we mean that the negation  $\neg\Phi$  holds everywhere on  $\underline{a}$ ). Moreover, one can decide between these two possibilities “syntactically”, just from the structure of  $\Phi$ , without any information about  $\underline{a}$  (except for a guarantee of the fast growth).

This will be expressed more precisely below, but first we formalize “growing sufficiently fast”.



**Definition 2.1.** Let  $R > 2$  be a real number. We call a sequence  $\underline{a} = (a_1, a_2, \dots, a_n)$   $R$ -growing if  $a_1 \geq R$  and  $a_{i+1} \geq a_i^R$ ,  $i = 1, 2, \dots, n-1$ .

**Observation 2.2.** For every 1-dimensional semialgebraic predicate  $\Phi = \Phi(x_1, \dots, x_k)$  there exists  $R > 1$  such that either  $\Phi$  holds everywhere on every  $R$ -growing sequence  $\underline{a}$ , or  $\Phi$  holds nowhere on every such  $\underline{a}$ . Moreover, it is easy to decide by inspection of  $\Phi$  which of these cases hold (without explicit knowledge of  $R$ ).

*Sketch of proof.* First we note that if  $p(x_1, \dots, x_k)$  is a polynomial and  $R$  is sufficiently large in terms of the degree and maximum absolute value of the coefficients of  $p$ , then the sign of  $p(a_{i_1}, a_{i_2}, \dots, a_{i_k})$ ,  $i_1 < \dots < i_k$ , is given by the sign of the coefficient of the lexicographically largest monomial present in  $p(x_1, \dots, x_k)$  (where we compare monomials first according to the power of  $x_k$ , then according to the power of  $x_{k-1}$ , etc.). The validity of  $\Phi$  can then be resolved based on the signs of the polynomials appearing in it.  $\square$

**A Ramsey-type result with  $R$ -growing sequences.** The next idea in our approach is that *every* sufficiently long sequence  $\underline{a}$  “contains” a long  $R$ -growing sequence, with a suitable meaning of “contains”.

Of course, not all long sequences contain  $R$ -growing subsequences; at the very least, we also have to consider reversals of  $R$ -growing sequences. Perhaps less obviously, we have sequences like

$$A, A+1, A+2, \dots, A+N,$$

where  $A$  is a number larger than  $N$ , which contain neither fast-growing subsequences nor their reversals.

In this case, we can find a *translate* of an  $R$ -growing sequence, of the form  $A + b_1, A + b_2, \dots, A + b_n$ , as a subsequence. For a slightly more sophisticated example, we consider the sequence

$$A + \frac{B}{1}, A + \frac{B}{2}, \dots, A + \frac{B}{N},$$

with two large parameters  $A, B$ . Here we have a subsequence obtained from an  $R$ -growing sequence  $(b_1, \dots, b_n)$  by the rational transformation  $x \mapsto A + \frac{B}{x}$ .

A key insight, coming from the thesis of Rosenthal [16], is that it suffices to consider rational transformations of simple form and with at most two parameters (like  $A$  and  $B$  above). To state this formally, we introduce the next definition.

**Definition 2.3.** A  $t$ -parametric transformation is a rational function  $f = f(x, X_1, \dots, X_t)$  in  $t+1$  variables.

Let  $\underline{b}$  be a sequence of length  $n$ , let  $f$  be a  $t$ -parametric transformation, and let  $A_1, \dots, A_t$  be real numbers. We write  $f(\underline{b}, A_1, \dots, A_t)$  for the sequence whose  $i$ th term is  $f(b_i, A_1, \dots, A_t)$  for all  $i = 1, 2, \dots, n$ .

If  $\underline{a}$  is a sequence of length  $N$ , we say that  $\underline{b}$  as above has a  $t$ -parametric embedding into  $\underline{a}$  via  $f$  if, for some  $A_1, \dots, A_t \in \mathbb{R}$ ,  $f(\underline{b}, A_1, \dots, A_t)$  is a subsequence of  $\underline{a}$ .

Finally, if  $\mathcal{F}$  is a set of  $t$ -parametric transformations, we say that  $\underline{b}$  has a  $t$ -parametric  $\mathcal{F}$ -embedding into  $\underline{a}$  if it has a  $t$ -parametric embedding into  $\underline{a}$  via some  $f \in \mathcal{F}$ .

Now we can state a key Ramsey-type result. Let  $\mathcal{F}_0$  stand for the set of the following two 2-parametric transformations:

$$f_1(x, X, Y) := X + Yx, \quad f_2(x, X, Y) := X + \frac{Y}{x}.$$

**Proposition 2.4.** *For every  $n$  and  $R$  there exists  $N$  such that for every sequence  $\underline{a}$  of length  $N$  there is an  $R$ -growing sequence  $\underline{b}$  of length  $n$  such that  $\underline{b}$  or  $\text{rev}(\underline{b})$ , the reversal of  $\underline{b}$ , has a 2-parametric  $\mathcal{F}_0$ -embedding into  $\underline{a}$ . Quantitatively, it suffices to take*

$$N := 2^{R^{2n}},$$

*provided that  $R \geq R_0$  for a sufficiently large constant  $R_0$ .*

The proof is presented in Section 4 below. A similar result, without the quantitative bound and stated in a different language, is implicit in Rosenthal [16].

### 3 Deciding predicates on transformed $R$ -growing sequences

Next, we would like to generalize Observation 2.2 from an  $R$ -growing sequence to a sequence  $\underline{c}$  obtained from an  $R$ -growing sequence  $\underline{b}$  (or its reversal) by a 2-parametric transformation.

Thus, let  $\underline{c} = f(\underline{b}, A, B)$ , for some 2-parametric transformation  $f(x, X, Y)$  and some  $A, B \in \mathbb{R}$ . For a given polynomial  $p(x_1, \dots, x_k)$ , we would like to understand the sign of  $p(c_{i_1}, \dots, c_{i_k})$ .

Let us substitute  $x_i = f(y_i, X, Y)$  into  $p(x_1, \dots, x_k)$ , and write the resulting rational function as the quotient of two polynomials (in  $y_1, \dots, y_k, X, Y$ ). We consider the sign of each them separately.

Thus, let  $q(y_1, \dots, y_k, X, Y)$  be one of these two polynomials; we write it as a polynomial in  $y_1, \dots, y_k$  whose coefficients are polynomials in  $X, Y$ :

$$q(y_1, \dots, y_k, X, Y) = \sum_{\alpha \in \Lambda} q_\alpha(X, Y) y_1^{\alpha_1} \cdots y_k^{\alpha_k},$$

where  $\Lambda = \Lambda(q)$  is the set of all multiindices  $\alpha = (\alpha_1, \dots, \alpha_k)$  such that the coefficient  $q_\alpha(X, Y)$  of  $y_1^{\alpha_1} \cdots y_k^{\alpha_k}$  is not identically zero.

Unlike in Observation 2.2, here we cannot in general determine the sign of  $q(b_1, \dots, b_k, A, B)$  just from the knowledge of the polynomial  $q$  and from the fact that  $\underline{b}$  is  $R$ -growing, since we do not know the signs and order of magnitude of the coefficients  $q_\alpha(A, B)$ .

Let us define, for  $\alpha, \beta \in \Lambda$ ,

$$\rho_{\alpha\beta} := \frac{q_\alpha(A, B)}{q_\beta(A, B)}$$

(where, for simpler notation, we put  $\rho_{\alpha\beta} := \infty$  if  $q_\beta(A, B) = 0$ ). We observe if each  $\rho_{\alpha\beta}$  is either “not too large” (considerably smaller than  $b_1$ ) or “very large” (much larger than  $b_k$ ), then the sign of  $q(b_1, \dots, b_k, A, B)$  can again be determined. The following definition captures this condition.

**Definition 3.1.** *Let  $q = q(y_1, \dots, y_k, X, Y)$  be a polynomial and let  $A, B \in \mathbb{R}$ . We say that an  $R$ -growing sequence  $\underline{b} = (b_1, \dots, b_n)$  is  $R$ -well-placed w.r.t.  $q, A, B$  if, for every  $\alpha, \beta \in \Lambda(q)$ , either*

- $\rho_{\alpha\beta}$  is dwarfed by  $\underline{b}$ , meaning that  $|\rho_{\alpha\beta}| \leq b_1/R$ , or
- $\rho_{\alpha\beta}$  is gigantic for  $\underline{b}$ , meaning that  $|\rho_{\alpha\beta}| \geq b_n^R$ .



In this situation we define the type of  $q$  w.r.t.  $A, B$ , and  $\underline{b}$  as the pair  $(\sigma, \tau)$ , where  $\sigma: \Lambda \rightarrow \{-1, 0, +1\}$  is given by  $\sigma(\alpha) := \text{sgn } q_\alpha(A, B)$ , and  $\tau: \Lambda^2 \rightarrow \{D, G\}$  is given by

$$\tau(\alpha, \beta) := \begin{cases} D & \text{if } \rho_{\alpha\beta} \text{ is dwarfed by } \underline{b}, \\ G & \text{if } \rho_{\alpha\beta} \text{ is gigantic for } \underline{b}. \end{cases}$$

We also want to extend these notions considerations from a single polynomial  $q$  to a collection of polynomials coming from a predicate (or set of predicates).

Let  $\Phi$  be a (finite) set of semialgebraic predicates and let  $f = f(x_1, \dots, x_k, X, Y)$  be a 2-parametric transformation. For every polynomial  $p(x_1, \dots, x_k)$  occurring in some predicate of  $\Phi$  we proceed as above, i.e., we perform the substitution  $x_i = f(y_1, \dots, y_k, X, Y)$ , and we express the resulting rational function as a quotient of two polynomials in  $y_1, \dots, y_k, X, Y$ . Let  $Q = Q(\Phi, f)$  stand for the collection of all the polynomials obtained in this way.

Then we will talk about an  $R$ -growing sequence  $\underline{b}$  being  $R$ -well-placed w.r.t.  $Q, A, B$  (meaning that it is  $R$ -well-placed w.r.t.  $q, A, B$  for every  $q \in Q$ ), and about the type of  $Q$  w.r.t.  $A, B$ , and  $\underline{b}$  (this is the  $|Q|$ -tuple  $((\sigma_q, \tau_q) : q \in Q)$ , where  $(\sigma_q, \tau_q)$  is the type of  $q$  w.r.t.  $A, B$ , and  $\underline{b}$ ).

Next, we extend our Ramsey-type result (Proposition 2.4) so that it yields well-placed sequences.

**Corollary 3.2.** *Let  $\mathcal{F}_0$  be as in Proposition 2.4, and let  $\Phi$  be a set of semialgebraic predicates. Then there is a constant  $C_1 = C_1(\Phi)$  such that for every  $n, R$  there exists*

$$N \leq 2^{R^{C_1 n}}$$

*such that for every sequence  $\underline{a}$  of length  $N$  there is an  $R$ -growing sequence  $\underline{b}$  of length  $n$  such that  $f(\underline{b}, A, B)$  or  $f(\text{rev}(\underline{b}), A, B)$  is a subsequence of  $\underline{a}$  for some  $f \in \mathcal{F}_0$  and  $A, B \in \mathbb{R}$ , and moreover,  $\underline{b}$  is  $R$ -well-placed w.r.t.  $Q(\Phi, f), A, B$ .*

*Proof.* Let us take  $n' := C(n+2)$ , where  $C$  is a sufficiently large number depending on  $\Phi$  and  $\mathcal{F}_0$ , and use Proposition 2.4 with  $n'$  instead of  $n$ . Given a sufficiently long sequence  $\underline{a}$ , we find an  $R$ -growing sequence  $\underline{b}$  of length  $n'$  as in the conclusion of Proposition 2.4, such that  $f(\underline{b}, A, B)$  or  $f(\text{rev}(\underline{b}), A, B)$  is a subsequence of  $\underline{a}$ .

Now  $\underline{b}$  is not necessarily  $R$ -well-placed w.r.t.  $Q(\Phi, f), A, B$ , and so we consider all the values  $\rho_{\alpha\beta}$  as in Definition 3.1, generated from all the polynomials in  $Q(\Phi, f)$ . These divide the real axis into intervals, whose number can be assumed to be at most  $C$ . Among these intervals, we fix one containing at least  $n'/C = n+2$  terms of  $\underline{b}$ . We take the (contiguous) subsequence of  $\underline{b}$  contained in this interval, and we delete the first and last terms. The remaining  $n$ -term subsequence is  $R$ -well-placed w.r.t.  $Q(\Phi, f), A, B$ .  $\square$

Here is an analog of Observation 2.2.

**Lemma 3.3.** *Let  $\Phi$  be a predicate, let  $\underline{b}$  be an  $R$ -growing sequence, and let  $\underline{c} = f(\underline{b}, A, B)$  for some 2-parametric transformation  $f$ . Suppose that  $\underline{b}$  is  $R$ -well-placed w.r.t.  $Q(\Phi, f), A, B$ , and that  $R$  is sufficiently large in terms of  $\Phi$  and  $f$ . Then  $\Phi$  holds either everywhere on  $\underline{c}$  or nowhere on it, and these possibilities can be distinguished based on  $\Phi, f$ , and the type of  $Q(\Phi, f)$  w.r.t.  $A, B, \underline{b}$  (without knowing  $A, B$  and  $\underline{b}$ ). A similar statement holds for the validity of  $\Phi$  on  $\text{rev}(\underline{c})$ .*

*Sketch of proof.* It suffices to check that if  $q = q(y_1, \dots, y_k, X, Y) \in Q(\Phi, f)$  is a single polynomial, then  $\text{sgn } q(b_{i_1}, \dots, b_{i_k}, A, B)$  is the same for all choices of  $i_1 < \dots < i_k$  (and it can be deduced from the type of  $q$ ).

The type of  $q$  w.r.t.  $A, B, \underline{b}$  gives us the signs of the terms  $q_\alpha(A, B)b_{i_1}^{\alpha_1} \dots b_{i_k}^{\alpha_k}$ ,  $\alpha \in \Lambda(q)$ . We just need to check that whenever  $\alpha, \beta \in \Lambda(q)$ ,  $\alpha \neq \beta$ , the absolute values of the terms  $q_\alpha(A, B)b_{i_1}^{\alpha_1} \dots b_{i_k}^{\alpha_k}$  and  $q_\beta(A, B)b_{i_1}^{\beta_1} \dots b_{i_k}^{\beta_k}$  have different orders of magnitude. If one of  $q_\alpha(A, B)$  and  $q_\beta(A, B)$  is 0, we are done. If they are both nonzero and  $\rho_{\alpha\beta}$  is gigantic, the first term wins; if  $\rho_{\beta\alpha}$  is gigantic, the second term wins; and if both  $\rho_{\alpha\beta}$  and  $\rho_{\beta\alpha}$  are dwarfed, then the comparison is lexicographic according to  $\alpha$  and  $\beta$ , as in Observation 2.2.  $\square$

*Proof of Proposition 1.6.* This proposition is an immediate consequence of Corollary 3.2 and Lemma 3.3.  $\square$

*Proof of Theorem 1.4.* Let  $k$  be the maximum arity of a predicate in  $\Phi$ , and let  $n_0 := \text{ES}_\Phi(k)$  (this is well defined since  $\Phi$  is Erdős–Szekeres).

Let us consider  $n \geq n_0$ , let  $N := 2^{R^{C_1 n}}$  be as in Corollary 3.2, and let  $\underline{a}$  be an arbitrary sequence of length  $N$ . Corollary 3.2 yields an  $R$ -growing sequence  $\underline{b}$  of length  $n$  such that, for some  $f \in \mathcal{F}_0$  and  $A, B \in \mathbb{R}$ ,  $\underline{c} := f(\underline{b}, A, B)$  or  $\underline{c} := f(\text{rev}(\underline{b}), A, B)$  is a subsequence of  $\underline{a}$ . Moreover,  $\underline{b}$  is  $R$ -well-placed w.r.t.  $Q(\Phi, f), A, B$ . Then by Lemma 3.3, each  $\Phi \in \Phi$  holds either everywhere on  $\underline{c}$  or nowhere on  $\underline{c}$ .

Since  $n \geq n_0 = \text{ES}_\Phi(k)$ , the sequence  $\underline{c}$  contains at least one  $k$ -tuple on which some  $\Phi \in \Phi$  holds, and hence this  $\Phi$  holds everywhere on  $\underline{c}$ . This shows that  $\text{ES}_\Phi(n) \leq 2^{R^{C_1 n}}$  for all  $n \geq n_0$ . Since  $R$  depends only on  $\Phi$ , this gives the desired bound  $\text{ES}_\Phi(n) \leq 2^{2^{O(n)}}$  for all  $n \geq n_0$ , and the finitely many  $n < n_0$  can be taken care of by setting  $C$  sufficiently large.  $\square$

The way we obtained the existence of  $n_0$  in the proof above was nonconstructive. Alternatively, one can obtain an explicit dependence of  $n_0$  on  $\Phi$  using the idea in Section 6.

## 4 Proof of the Ramsey-type result for sequences

We begin with a quantitative version of the “dichotomy lemma” of Rosenthal [16]. Let us say that a sequence  $\underline{b} = (b_1, \dots, b_n)$  satisfies the *doubling differences condition*, or the DDC for short, if  $|b_k - b_i| \geq 2|b_j - b_i|$  holds for every  $i < j < k$  (or, in our previous terminology, the predicate “ $|x_3 - x_1| \geq 2|x_2 - x_1|$ ” holds everywhere on  $\underline{b}$ ). We note that the differences in a sequence  $\underline{b}$  satisfying the DDC grow (at least) exponentially, e.g.,  $|b_i - b_1| \geq 2^{i-1}|b_2 - b_1|$ .

**Lemma 4.1.** *Let  $n$  be a natural number, and let  $N = \binom{2n}{n} \leq 4^n$ . Then every (strictly) increasing sequence  $\underline{a}$  of real numbers of length  $N$  has a subsequence  $\underline{b}$  of length  $n$  such that one of the two (increasing) sequences  $\underline{b}$  and  $\text{rev}(-\underline{b})$  (where  $-\underline{b}$  stands for  $(-b_1, \dots, -b_n)$ ) satisfies the DDC.*

*Proof.* Before proceeding, let us note that if we are not interested in the quantitative bound, the existence of a suitable  $N$  follows easily from Ramsey’s theorem for triples. Indeed, we observe that, assuming  $b_i < b_j < b_k$ , the DDC  $b_k - b_i \geq 2(b_j - b_i)$  is equivalent to  $b_k - b_j \geq b_j - b_i$ . Given an increasing sequence  $(a_1, a_2, \dots, a_N)$ , we color a triple  $\{i, j, k\} \subseteq [N]$ ,  $i < j < k$ , red if  $a_k - a_j \geq a_j - a_i$ , and blue otherwise. Now a red homogeneous subset corresponds to a subsequence  $\underline{b}$  satisfying the DDC, and a blue homogeneous subset corresponds to a

subsequence  $\underline{b}$  with  $\text{rev}(\underline{b})$  satisfying the DDC. The latter sequence is decreasing rather than increasing, but we can repair this by considering  $\text{rev}(-\underline{b})$ , since negation preserves the DDC.

Now we present the proof giving a better quantitative bound; it resembles one of the well-known proofs of the Ramsey theorem for graphs. Let us define  $N(k, \ell)$  as the smallest  $N$  such that every increasing sequence  $\underline{a}$  of length  $N$  contains a subsequence  $\underline{b}$  of length  $k$  satisfying the DDC, or a subsequence  $\underline{b}$  of length  $\ell$  with  $\text{rev}(\underline{b})$  satisfying the DDC.

We have the initial conditions  $N(2, \ell) = 2$  and  $N(k, 2) = 2$ , and below we will derive the recurrence

$$N(k, \ell) \leq N(k-1, \ell) + N(k, \ell-1) - 1, \quad k, \ell \geq 3.$$

It is well known, and easy to check, that this implies  $N(k, \ell) \leq \binom{k+\ell}{k}$ , and so the bound in the lemma follows.

To verify the recurrence, let  $N := N(k-1, \ell) + N(k, \ell-1) - 1$  and let  $\underline{a}$  be a nondecreasing sequence of length  $N$ . We divide the interval  $[a_1, a_N]$  into the left and right subinterval by the midpoint  $\frac{1}{2}(a_1 + a_N)$ , and we let  $\underline{a}'$  and  $\underline{a}''$  be the parts of  $\underline{a}$  in the left and right subinterval, respectively. Then  $\text{len } \underline{a}' \geq N(k-1, \ell)$  or  $\text{len } \underline{a}'' \geq N(k, \ell-1)$ .

In the former case, if  $\underline{a}'$  contains a subsequence  $\underline{b}$  of length  $\ell$  with  $\text{rev}(\underline{b})$  satisfying the DDC, we are done. Otherwise,  $\underline{a}'$  has a subsequence  $\underline{b}'$  of length  $k-1$  satisfying the DDC, and it is easy to see that the sequence obtained by appending  $a_N$  to  $\underline{b}'$  also satisfies the DDC. The case of  $\text{len } \underline{a}'' \geq N(k, \ell-1)$  is analogous. This concludes the proof.  $\square$

Rather than using Lemma 4.1 directly, we will apply the following simple consequence, where the DDC is strengthened to  $R$ -fold expansion of the differences.

**Corollary 4.2.** *Let  $n$  be a natural number, let  $R > 1$ , and set  $r := \lceil \log_2 R \rceil$ . Then there is  $N \leq 4^{r(n-1)+1}$  such that every (strictly) increasing sequence  $\underline{a}$  of real numbers of length  $N$  has a subsequence  $\underline{b}$  of length  $n$  such that the predicate “ $x_3 - x_1 \geq R(x_2 - x_1)$ ” holds everywhere on one of the two (increasing) sequences  $\underline{b}$  and  $\text{rev}(-\underline{b})$ .*

*Proof.* Select a subsequence  $\underline{b}'$  of  $\underline{a}$  of length  $r(n-1) + 1$  as in Lemma 4.1, and define  $b_i := b'_{r(i-1)+1}$ ,  $i = 1, 2, \dots, n$ .  $\square$

*Proof of Proposition 2.4.* Let  $n$  and  $R$  be given, let  $N$  be sufficiently large as in the proposition, and let  $\underline{a}$  be a sequence of length  $N$ .

First, if some number occurs at least  $n$  times in  $\underline{a}$ , we are done, since the 1-parametric transformation  $f(x, X) = X$  (a special case of both  $f_1$  and  $f_2$  with  $Y = 0$ ) embeds any  $n$ -term sequence into  $\underline{a}$ . Otherwise, we may pass to a subsequence  $\underline{a}'$  with all terms distinct and of length at least  $N/n$ . Next, by the Erdős-Szekeres lemma, we can further pass to a subsequence  $\underline{a}''$ , with  $\text{len}(\underline{a}'') \geq \sqrt{N/n}$ , that is either increasing or decreasing. We consider only the increasing case, the decreasing one being symmetric.

Next, applying Corollary 4.2, we obtain a subsequence  $\underline{a}^{(3)}$  of  $\underline{a}''$ , where  $\text{len } \underline{a}^{(3)} \leq 4^{r(\text{len } \underline{a}^{(3)}-1)+1}$ , such that “ $x_3 - x_1 \geq R(x_2 - x_1)$ ” holds everywhere on  $\underline{a}^{(3)}$  or  $\text{rev}(-\underline{a}^{(3)})$ . Let us again discuss only the former case.

In this case we form a new sequence  $\underline{b}^{(3)}$  of length  $\text{len } \underline{a}^{(3)} - 1$  by setting  $b_i^{(3)} := a_{i+1}^{(3)} - a_1^{(3)}$ . We note that  $\underline{b}^{(3)}$  satisfies  $b_1^{(3)} > 0$  and  $b_{i+1}^{(3)} \geq R b_i^{(3)}$ , and it embeds into  $\underline{a}$  via the 1-parametric transformation  $f(x, X) = x + X$  for  $X := a_1^{(3)}$ .

Now we apply Corollary 4.2 again, this time to the sequence  $\underline{\ell}$  with  $\ell_i = \log b_i^{(3)}$ , and this yields a subsequence  $\underline{b}^{(4)}$  of  $\underline{b}^{(3)}$ , such that

$$\frac{b_k^{(4)}}{b_i^{(4)}} \geq \left( \frac{b_j^{(4)}}{b_i^{(4)}} \right)^R$$

holds for every  $i < j < k$ , or a similar relation holds for  $\text{rev}(1/\underline{b}^{(4)})$ . As expected, we again deal explicitly only with the first possibility, and we set  $c_i^{(4)} := b_{i+1}^{(4)}/b_1^{(4)}$ ,  $i = 1, 2, \dots, \text{len}(\underline{b}^{(4)}) - 1$ .

By this choice, we have  $c_{i+1}^{(4)} \geq (c_i^{(4)})^R$ . Moreover, since  $b_{i+1}^{(4)} \geq Rb_i^{(4)}$  for all  $i$ , we also obtain  $c_1^{(4)} = b_2^{(4)}/b_1^{(4)} \geq R$ , and so  $\underline{c}^{(4)}$  is  $R$ -growing. It is also clear that  $\underline{c}^{(4)}$  embeds into  $\underline{b}^{(3)}$  via the 1-parametric transformation  $g(x, X) := x/X$  with  $X := b_1^{(4)}$ , and hence it embeds into  $\underline{a}$  via the 2-parametric transformation  $h(x, X, Y) = f(g(x, X), Y)$ . Thus,  $\underline{c}^{(4)}$  is our desired sequence (called  $\underline{b}$  in the proposition).

By following the chain of length estimates backwards, we get that

$$N \geq n4^{2(r(4^{r(n-1)+2}-1)+2)}$$

suffices for the whole argument. Assuming that  $R$  exceeds a suitable constant (not very large, actually), a series of rough estimates shows that the right-hand side can be bounded by  $2^{R^{2n}}$  as claimed.  $\square$

## 5 The decision algorithm

Before formulating the algorithm for deciding whether a given set of predicates  $\Phi$  is Erdős–Szekeres, we introduce some additional terminology.

If  $q = q(y_1, \dots, y_k, X, Y)$  is a polynomial, a *candidate type* for  $q$  is a pair  $(\sigma, \tau)$  of arbitrary functions  $\sigma: \Lambda \rightarrow \{-1, 0, +1\}$  and  $\tau: \Lambda^2 \rightarrow \{D, G\}$ , where  $\Lambda$  is as introduced above in Definition 3.1. A *candidate type* for a collection  $Q$  of polynomials (all in the variables  $y_1, \dots, y_k, X, Y$ ) is an arbitrary sequence  $T = ((\sigma_q, \tau_q) : q \in Q)$ , where each  $(\sigma_q, \tau_q)$  is a candidate type for  $q$ .

**Definition 5.1.** *We call a candidate type  $T$  for a collection  $Q$  of polynomials feasible if for every  $R$  and every  $n$  there exist  $A, B \in \mathbb{R}$  and an  $R$ -growing sequence  $\underline{b}$  of length  $n$  that is  $R$ -well-placed w.r.t.  $Q$ , such that the type of  $Q$  w.r.t.  $A, B$ , and  $\underline{b}$  equals  $T$ . Any such sequence  $\underline{b}$  is called a feasible  $R$ -growing sequence for  $Q$  and  $T$ .*

We note that our definition of feasibility does not use any particular value of  $R$ , but has the quantification “for all  $R$ ”. Accordingly, we will never use an explicit value of  $R$  in our algorithm.

In Section 6 below, we will provide a subroutine that tests feasibility of a given candidate type for a given collection  $Q$  of polynomials. Now we present the main algorithm, in which we use this feasibility testing as a black box.

### Algorithm for testing if $\Phi$ is Erdős–Szekeres.

*Input:* A finite set  $\Phi$  of semialgebraic predicates.

*Output:* YES if  $\Phi$  is Erdős–Szekeres, NO otherwise.

1. Let  $\mathcal{F}_0$  be the set of two 2-parametric transformations as in Proposition 2.4. For each  $f$  let  $Q(\Phi, f)$  be the collection of polynomials obtained from  $\Phi$  (as introduced after Definition 3.1). Perform the following steps for every  $f \in \mathcal{F}_0$  and every candidate type  $T$  for  $Q(\Phi, f)$ . If all these steps are completed without returning NO, return YES and finish.
2. Test the feasibility of  $T$  (as specified in Section 6 below). If  $T$  is feasible, continue with the next step; otherwise, continue at Step 1 with the next  $T$  or next  $f$ .
3. Determine, by the method of Lemma 3.3, whether all  $\neg\Phi$  for  $\Phi \in \Phi$ , hold everywhere on  $f(\underline{b}, A, B)$  or on  $f(\text{rev}(\underline{b}), A, B)$ , where  $\underline{b}$  is a feasible  $R$ -growing sequence for  $Q(\Phi, f)$  and  $T$  (here  $A, B$  are the corresponding parameter values, which we need not determine explicitly; similarly,  $R$  is only assumed to be sufficiently large). If they do, return NO and finish the whole algorithm. If not, continue at Step 1 with the next  $T$  or next  $f$ .

*Proof of Theorem 1.7 (assuming the feasibility testing).* The algorithm is clearly finite, so it suffices to verify that its answer is always correct. With the tools developed above, the proof is routine.

First suppose that the algorithm returns NO; then this is because in Step 3 it has exhibited arbitrarily long sequences  $\underline{c}$  on which all  $\neg\Phi$ ,  $\Phi \in \Phi$ , hold everywhere. So  $\Phi$  is certainly not Erdős–Szekeres.

Now suppose that  $\Phi$  is not Erdős–Szekeres; this means that for some  $n_0$ , there exist arbitrarily long sequences for which no  $\Phi \in \Phi$  holds everywhere on any subsequence of length  $n_0$ . By Ramsey’s theorem, this means that there are arbitrarily long sequences  $\underline{a}$  on which each of  $\neg\Phi$ ,  $\Phi \in \Phi$ , holds everywhere. (Alternatively, we can avoid using Ramsey’s theorem and argue as in the proof of Theorem 1.4 at the end of Section 3.) For every  $N$ , fix such a sequence  $\underline{a}^{(N)}$  of length  $N$ .

For every  $n$ , by Corollary 3.2, there is  $N = N(n)$  such that for some  $n$ -growing  $n$ -term sequence  $\underline{b}^{(n)}$ , some  $f^{(n)} \in \mathcal{F}_0$ , and some  $A^{(n)}, B^{(n)} \in \mathbb{R}$ , one of the sequences  $f(\underline{b}^{(n)}, A^{(n)}, B^{(n)})$  and  $f(\text{rev}(\underline{b}^{(n)}), A^{(n)}, B^{(n)})$  is a subsequence of  $\underline{a}^{(N(n))}$ , and  $\underline{b}^{(n)}$  is  $n$ -well-placed w.r.t.  $Q(\Phi, f^{(n)})$ ,  $A^{(n)}, B^{(n)}$ . Let  $T^{(n)}$  be the type of  $Q(\Phi, f^{(n)})$  w.r.t.  $A^{(n)}, B^{(n)}$ , and  $\underline{b}^{(n)}$ .

Since  $\mathcal{F}_0$  is finite and there are finitely many types, there is an infinite subsequence  $n_1 < n_2 < \dots$  such that all the  $f^{(n_j)}$  are equal to the same  $f$  and all the  $T^{(n_j)}$  are equal to the same  $T$ . But then we get that  $T$ , regarded as a candidate type for  $Q(\Phi, f)$ , is feasible, and Step 3 returns NO for this  $f$  and  $T$ . This concludes the proof.  $\square$

## 6 Testing the feasibility of a candidate type

The problem of testing feasibility of a given candidate type for a collection of polynomials can be recast in the following (slightly more general) terms. We are given a finite collection  $\mathcal{Q}$  of bivariate polynomials (the  $q_\alpha(X, Y)$  in the setting of Definition 5.1), and for each  $q \in \mathcal{Q}$ , a sign  $\sigma_q \in \{-1, 0, +1\}$  is specified. Moreover, we are given two finite collections  $\mathcal{D}$  and  $\mathcal{G}$  of rational functions in the variables  $X, Y$  (in the setting of Definition 5.1,  $\mathcal{D}$  consists of the  $\rho_{\alpha\beta}$  that should be dwarfed by the desired feasible sequence, while  $\mathcal{G}$  are the  $\rho_{\alpha\beta}$  that should be gigantic for it). The numerators and denominators of the rational functions in  $\mathcal{D}$  and in  $\mathcal{G}$  belong to  $\mathcal{Q}$ , and thus their signs are under control. Moreover, we may assume that the prescribed signs for all the denominators are nonzero.

The question then is whether we can make a large enough gap between the values of the functions in  $\mathcal{D}$  and those in  $\mathcal{G}$  so that an  $n$ -term  $R$ -growing sequence fits there. Formally, we thus ask for the validity of the formula

$$\Psi := \forall n \in \mathbb{N} \forall R \exists L, H : (L \geq R) \wedge (H \geq L^{R^{n+2}}) \wedge \Xi(L, H),$$

where

$$\begin{aligned} \Xi(L, H) := \exists X, Y : & \left( \bigwedge_{q \in \mathcal{Q}} \operatorname{sgn} q(X, Y) = \sigma_q \right) \\ & \wedge \left( \bigwedge_{\rho \in \mathcal{D}} \rho(X, Y) \leq L \right) \wedge \left( \bigwedge_{\rho \in \mathcal{G}} \rho(X, Y) \geq H \right). \end{aligned}$$

We would like to test the validity of  $\Psi$  using a decision algorithm for the first-order theory of real-closed fields (the first such algorithm is due to Tarski [18], and we refer, e.g., to [2] to more recent ones). While  $\Xi(L, H)$  can easily be rewritten to a first-order formula in the theory of the reals (since  $\operatorname{sgn} q(X, Y) = \sigma_q$  is expressed as a polynomial inequality, and the inequalities  $\rho(X, Y) \leq L$  or  $\rho(X, Y) \geq H$  can be multiplied by the denominator, since we know its sign),  $\Psi$  is not such a formula: quantification over the natural numbers, and more importantly, the exponential function, do not belong to the first-order theory of the reals.

To remedy this, we replace  $\Psi$  by the formula

$$\Psi^* := \forall R \exists L : (L \geq R) \wedge \forall H \Xi(L, H).$$

The validity of  $\Psi$  can be tested by the decision algorithms mentioned above, and so for completing our subroutine for feasibility testing, it suffices to prove the following.

**Lemma 6.1.** *The formulas  $\Psi$  and  $\Psi^*$  are equivalent.*

*Proof.* Clearly  $\Psi^*$  implies  $\Psi$ . For the reverse implication, let us consider the function

$$h(L) := \sup\{H : \Xi(L, H)\}.$$

Supposing that  $\Psi^*$  does not hold, we see that there are arbitrarily large values of  $L$  such that  $h(L)$  is finite. Since  $h$  is a semialgebraic function, it has to be finite for all  $L \geq L_0$  for some  $L_0$ . But if  $\Psi$  did hold, then for every  $m$  there are arbitrarily large values of  $L$  with  $h(L) \geq L^m$ , while it is easy to check, and well known, that a semialgebraic function may have at most a polynomial growth. Hence  $\Psi$  does not hold and the lemma is proved.  $\square$

## 7 A lower bound for the Ramsey function

*Proof of Proposition 1.5.* We recall that the *cross-ratio* of an ordered 4-tuple  $(z_1, \dots, z_4)$  of real numbers is defined as

$$(z_1, z_2; z_3, z_4) := \frac{(z_1 - z_3)(z_2 - z_4)}{(z_2 - z_3)(z_1 - z_4)}.$$

As is well known, the cross-ratio is invariant under projective transforms of the form  $x \mapsto (ax + b)/(cx + d)$  (assuming  $ad - bc \neq 0$ ), and in particular, it is invariant under the two



transformations  $f_1(x, X, Y) = X + Yx$  and  $f_2(x, X, Y) = X + Y/x$  of  $\mathcal{F}_0$  (unless  $Y = 0$ , that is).

We define the set of predicates  $\Phi = \{\Phi_1, \Phi_2, \Phi_3\}$  as follows:

$$\begin{aligned}\Phi_1(x_1, x_2) &:= x_1 = x_2, \\ \Phi_2(x_1, \dots, x_5) &:= \text{DISTINCT}(x_1, \dots, x_5) \wedge (x_1, x_2; x_3, x_4)^2 \geq 4 \\ &\quad \wedge (x_1, x_2; x_3, x_5)^2 \geq (x_1, x_2; x_3, x_4)^4, \\ \Phi_3(x_1, \dots, x_5) &:= \Phi_2(x_5, x_4, \dots, x_1).\end{aligned}$$

Here  $\text{DISTINCT}(x_1, \dots, x_5)$  has the expected meaning. The condition  $(x_1, x_2; x_3, x_4)^2 \geq 4$  should be read as  $|(x_1, x_2; x_3, x_4)| \geq 2$ , but it is written in squared form, so that it can be expressed as a polynomial inequality and we need not worry about the sign. A similar comment applies to the last condition of  $\Phi_2$ .

First we check that  $\Phi$  is Erdős–Szekeres. Thus, let  $n$  be given, let  $R$  be a suitable constant, and let  $\underline{a}$  be a sufficiently long sequence. By Proposition 2.4, there is an  $n$ -term  $R$ -growing sequence  $\underline{b}$  such that  $\underline{b}$  or  $\text{rev}(\underline{b})$  has a 2-parametric embedding in  $\underline{a}$  via  $\mathcal{F}_0$ .

It is easy to check that  $\Phi_2$  holds everywhere on any  $R$ -growing sequence with a sufficiently large constant  $R$  (intuitively it is clear that the cross-ratios have to grow fast, and formally it can be checked as in Observation 2.2), and consequently,  $\Phi_3$  holds everywhere on the reversal of an  $R$ -growing sequence.

Since  $\Phi_2$  and  $\Phi_3$  are expressed in terms of cross-ratios, and these are invariant under the transformations of  $\mathcal{F}_0$ , we get that if  $\underline{b}$   $\mathcal{F}_0$ -embeds into  $\underline{a}$ , then either the image of  $\underline{b}$  is a constant sequence, or then  $\Phi_2$  holds everywhere on the image of  $\underline{b}$ . Similarly, if  $\text{rev}(\underline{b})$   $\mathcal{F}_0$ -embeds into  $\underline{a}$ , then either the corresponding subsequence of  $\underline{a}$  is constant, or  $\Phi_3$  holds everywhere on it. Thus,  $\Phi$  is indeed Erdős–Szekeres.

Next, we want to prove a doubly exponential lower bound for  $\text{ES}_\Phi(n)$ . To this end, let  $\underline{a} = (1, 2, \dots, N)$ , and let  $\underline{c}$  be an  $n$ -term subsequence on which one of  $\Phi_2, \Phi_3$  holds everywhere ( $\Phi_1$  is out since the terms of  $\underline{a}$  are all distinct). By possibly replacing  $\underline{c}$  by its reversal, we may assume that  $\Phi_2$  holds everywhere on  $\underline{c}$ . Then the cross-ratios in  $\underline{c}$  satisfy  $|(c_1, c_2; c_3, c_4)| \geq 2$  and  $|(c_1, c_2; c_3, c_{i+1})| \geq |(c_1, c_2; c_3, c_i)|^2$ , and so  $|(c_1, c_2; c_3, c_n)| \geq 2^{2^{n-4}}$ . Since all terms of  $\underline{c}$  are integers, at least one of them has to be at least  $2^{2^{cn}}$ .  $\square$

## 8 Replacing a set of predicates by a single predicate

Here we consider two simple constructions for replacing a finite set  $\Phi = \{\Phi_1, \dots, \Phi_r\}$  of  $d$ -dimensional  $k$ -ary predicates with a single predicate  $\bar{\Phi}$ , in such a way that  $\Phi$  is Erdős–Szekeres iff  $\bar{\Phi}$  is.

In the first construction, we simply set  $\bar{\Phi} := \bigvee_{i=1}^r \Phi_i$ . Then if some  $\Phi_i$  holds everywhere on some sequence  $\underline{a}$ , then so does  $\bar{\Phi}$ . Conversely, assuming that  $\bar{\Phi}$  holds everywhere on a sequence  $\underline{a}$  of length  $N$ , we color each  $k$ -tuple  $j_1 < \dots < j_k$  of indices by a color  $i$  such that  $\Phi_i(\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_k})$  holds. Then by Ramsey’s theorem, some  $\Phi_i$  holds everywhere on a suitable subsequence of length  $n$ , provided that  $N$  is sufficiently large in terms of  $n, k$ , and  $r$ .

However, for this construction, it might happen that  $\text{ES}_{\{\bar{\Phi}\}}$  is much larger than  $\text{ES}_\Phi$ . Here is the second construction, which preserves the Ramsey function but increases the number of variables and the complexity of the predicate considerably.

**Lemma 8.1.** *Let  $\Phi = \{\Phi_1, \dots, \Phi_r\}$  be a set of  $k$ -ary predicates. Then there is a predicate  $\bar{\Phi}$  in  $rk$  variables such that  $\{\bar{\Phi}\}$  is Erdős–Szekeres iff  $\Phi$  is, and if they are Erdős–Szekeres, then  $\text{ES}_{\bar{\Phi}}(n) = \text{ES}_{\Phi}(n)$  for all  $n \geq rk$ .*

*Proof.* We first define  $(rk)$ -ary predicates  $\Psi_i$ ,  $i = 1, 2, \dots, r$ , where  $\Psi_i(\mathbf{x}_1, \dots, \mathbf{x}_{rk})$  expresses that  $\Phi_i$  holds everywhere on the sequence  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{rk})$ . Explicitly,

$$\Psi_i(\mathbf{x}_1, \dots, \mathbf{x}_{rk}) := \bigwedge_{1 \leq j_1 < \dots < j_k \leq rk} \Phi_i(\mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_k}).$$

Then we set  $\bar{\Phi}(\mathbf{x}_1, \dots, \mathbf{x}_{rk}) := \bigvee_{i=1}^r \Psi_i(\mathbf{x}_1, \dots, \mathbf{x}_{rk})$ .

Clearly, if some  $\Phi_i$  holds everywhere on a sequence  $\underline{a}$ , then so does  $\bar{\Phi}$ . Conversely, suppose that  $\bar{\Phi}$  holds everywhere on a sequence  $\underline{a}$  of length  $n \geq rk$ ; we claim that then some  $\Phi_i$  holds everywhere on  $\underline{a}$  as well.

If it were not the case, we fix, for every  $i = 1, 2, \dots, r$ , a  $k$ -tuple  $J_i \subseteq [n]$  such that  $\Phi_i$  does not hold on the corresponding  $k$  terms of  $\underline{a}$ . We consider the union  $\bigcup_{i=1}^r J_i$  and, if it has fewer than  $rk$  elements, we add the appropriate number of other elements of  $[n]$  (chosen arbitrarily) to it, obtaining an  $rk$ -element set  $J$ . Then  $\bar{\Phi}$  does not hold on the subsequence of  $\underline{a}$  indexed by  $J$ ; the resulting contradiction proves the lemma.  $\square$

## 9 Algebraic predicates, and the multipartite setting

In this section we prove the effective compactness for  $d$ -dimensional algebraic predicates (Theorem 1.8), as well as the effective compactness for semialgebraic predicates in the multipartite setting (Theorem 1.9). We begin with the multipartite setting, since the proof is formally somewhat simpler.

We will repeatedly use the following straightforward fact: for every set  $P \subseteq \mathbb{R}[x_1, \dots, x_t]$  of  $t$ -variate polynomials, each of degree at most  $D$ , there is a subset  $P_0 \subseteq P$  of at most  $\binom{D+t}{t}$  polynomials that defines the same variety in  $\mathbb{R}^t$  as  $P$ , where the variety defined by  $P$  is  $V(P) = \{\mathbf{x} \in \mathbb{R}^t : p(\mathbf{x}) = 0 \text{ for all } p \in P\}$ . Indeed, the vector space of all  $t$ -variate polynomials of degree at most  $D$  has dimension  $\binom{D+t}{t}$ , with the set of all monomials forming a basis, and if we choose  $P_0$  as a basis of the subspace generated by  $P$ , then we have  $V(P_0) = V(P)$ . We will refer to this fact as the *bounded-dimension argument*.

*Proof of Theorem 1.9.* Throughout the proof, by saying that a certain quantity is *bounded* we mean that it can be bounded from above by some explicit function of  $d$  (the space dimension),  $k$  (the arity of the predicate), and  $D$  (the maximum degree of the polynomials in the predicate).

To explain the idea of the proof in a simpler setting, we first assume that  $\Phi$  is *algebraic* and binary, i.e.,  $k = 2$ .

We need to prove only the implication (i)  $\Rightarrow$  (ii). So we assume that  $\Phi(\mathbf{x}_1, \mathbf{x}_2)$  holds everywhere on  $A_1 \times A_2$ , where  $|A_1| = |A_2| = N$  is large but so far unspecified. Our goal is to construct infinite sets  $X_1, X_2 \subseteq \mathbb{R}^d$  such that  $\Phi$  holds everywhere on  $X_1 \times X_2$ .

By the assumptions of the theorem,  $\Phi$  is a conjunction of polynomial equations of degree at most  $D$ , and by the bounded-dimension argument, we can assume that the number of equations is bounded.

The plan is the following: We are going to define two semialgebraic sets (and actually, algebraic varieties)  $V_{1,2}, V_{2,1} \subseteq \mathbb{R}^d$ , such that  $A_1 \subseteq V_{1,2}$ ,  $A_2 \subseteq V_{2,1}$ , and  $\Phi$  is easily checked

to hold on  $V_{1,2} \times V_{2,1}$ . We will also show that  $V_{1,2}$  and  $V_{2,1}$  can be defined by bounded-size formulas (i.e., the formula is a Boolean combination of at most  $m$  polynomial equations and inequalities, each of degree at most  $\bar{D}$ , where  $m$  and  $\bar{D}$  are bounded). Then we will use a result stating that the number of connected components of a semialgebraic set definable by a bounded-size formula is bounded (a specific result of this kind is a theorem of Warren [21, Theorem 2], based on the well-known Oleinik–Petrovskii–Milnor–Thom theorem [14, 11, 19] on the sum of Betti numbers of a real algebraic variety; also see [2, Chap. 7]). Finally, if  $N = |A_1|$  is larger than the number of components of  $V_{1,2}$ , then two points of  $A_1$  must be connected by a path, and hence  $V_{1,2}$  is infinite (and similarly for  $V_{2,1}$ ).

It remains to define  $V_{1,2}$  and  $V_{2,1}$  and to verify the claimed properties. Let  $F_1 = F_1(\mathbf{x}_1)$  be the formula  $\bigwedge_{\mathbf{a}_2 \in A_2} \Phi(\mathbf{x}_1, \mathbf{a}_2)$ , and let  $V_1 \subseteq \mathbb{R}^d$  be the variety  $\{\mathbf{x}_1 \in \mathbb{R}^d : F_1(\mathbf{x}_1)\}$ . We define  $F_2(\mathbf{x}_2) := \bigwedge_{\mathbf{a}_1 \in A_1} \Phi(\mathbf{a}_1, \mathbf{x}_2)$  and  $V_2$  similarly. We have  $A_i \subseteq V_i$ ,  $i = 1, 2$ , by the assumption.

We set

$$V_{1,2} := \left\{ \mathbf{x}_1 \in V_1 : (\forall \mathbf{x}_2 \in V_2) \Phi(\mathbf{x}_1, \mathbf{x}_2) \right\}$$

and

$$V_{2,1} := \left\{ \mathbf{x}_2 \in V_2 : (\forall \mathbf{x}_1 \in V_1) \Phi(\mathbf{x}_1, \mathbf{x}_2) \right\}.$$

By definition,  $\Phi$  holds everywhere on  $V_{1,2} \times V_2$ , as well as on  $V_1 \times V_{2,1}$ . Since  $V_{1,2} \subseteq V_1$  and  $V_{2,1} \subseteq V_2$ , we get that  $\Phi$  holds everywhere on  $V_{1,2} \times V_{2,1}$ .

It is easy to see that  $A_1 \subseteq V_{1,2}$  and  $A_2 \subseteq V_{2,1}$ , and we now want to argue that  $V_{1,2}$  and  $V_{2,1}$  can be defined by bounded-size formulas.

First, we  $V_1$  is the variety defined by the polynomials  $f(\mathbf{x}_1, \mathbf{a}_2)$ , where  $f$  is one of the polynomials in  $\Phi$  and  $\mathbf{a}_2 \in A_2$ . By the bounded-dimension argument, only a bounded number of these polynomials suffice to define  $V_1$ , and hence we can obtain a bounded-size formula  $\tilde{F}_1$  that is equivalent to  $F_1$ . Similarly, from  $F_2$  we obtain a bounded-size equivalent formula  $\tilde{F}_2$ .

Now we consider the formula  $F_{1,2}(\mathbf{x}_1)$  defining  $V_{1,2}$ . By definition,

$$F_{1,2}(\mathbf{x}_1) := F_1(\mathbf{x}_1) \wedge \forall \mathbf{x}_2 (F_2(\mathbf{x}_2) \Rightarrow \Phi(\mathbf{x}_1, \mathbf{x}_2)).$$

We now replace  $F_1$  and  $F_2$  with  $\tilde{F}_1$  and  $\tilde{F}_2$ , which gives a bounded-size formula describing  $V_{1,2}$ , and then we apply quantifier elimination to obtain a quantifier-free formula  $\bar{F}_{1,2}(\mathbf{x}_1)$  that is equivalent to  $F_{1,2}(\mathbf{x}_1)$  and thus also defines  $V_{1,2}$ . Moreover, by the properties of available quantifier-elimination methods for the first-order theory of the reals (see [2, Sec. 11.3]), the size of  $\bar{F}_{1,2}$  is still bounded. Hence we can bound the number of connected components of  $V_{1,2}$  as in the plan above, and the considered special case of Theorem 1.9 is proved.

Next, we still assume that  $\Phi = \Phi(\mathbf{x}_1, \dots, \mathbf{x}_k)$  is algebraic, but it can be  $k$ -ary for any  $k$ . The basic idea is the same as before, but this time we define a more complicated sequence of sets  $V_{i,J}$ , with  $i \in [k]$  and  $J \subseteq [k] \setminus \{i\}$ , inductively by

$$V_{i,\emptyset} := \left\{ \mathbf{x}_i \in \mathbb{R}^d : (\forall \mathbf{a}_1 \in A_1) \cdots (\forall \mathbf{a}_{i-1} \in A_{i-1}) (\forall \mathbf{a}_{i+1} \in A_{i+1}) \cdots (\forall \mathbf{a}_k \in A_k) \right. \\ \left. \Phi(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{x}_i, \mathbf{a}_{i+1}, \dots, \mathbf{a}_k) \right\}$$

and, for  $J \neq \emptyset$ ,

$$V_{i,J} := \left\{ \mathbf{x}_i \in \bigcap_{J' \subset J} V_{i,J'} : (\forall \mathbf{x}_j \in V_{j,J \setminus \{j\}}, j \in J) (\forall \mathbf{a}_j \in A_j, j \in [k] \setminus \{i\} \setminus J) \right. \\ \left. \Phi(\mathbf{x}_i; \mathbf{x}_j, j \in J; \mathbf{a}_j, j \in [k] \setminus \{i\} \setminus J) \right\}.$$

where the sequence of arguments of  $\Phi$  means that the  $i$ th argument equals  $\mathbf{x}_i$ , the  $j$ th argument equals  $\mathbf{x}_j$  for all  $j \in J$ , and the  $j$ th argument equals  $\mathbf{a}_j$  all  $j \in [k] \setminus \{i\} \setminus J$ .

By definition, for  $J' \subseteq J$  we have  $V_{i,J} \subseteq V_{i,J'}$ , and this then shows that  $\Phi$  holds everywhere on  $V_{1,[k] \setminus \{1\}} \times \cdots \times V_{k,[k] \setminus \{k\}}$ . Inductively it is also easy to show that  $A_i \subseteq V_{i,J}$  for all  $J$ .

It remains to show that each  $V_{i,J}$  can be described by a bounded-size formula, possibly involving quantifiers (then we apply quantifier elimination and bound the number of connected components as in the case  $k = 2$ ). We are going to show this by induction on  $|J|$ .

Thus, the inductive hypothesis is that we have bounded-size formulas  $\tilde{F}_{i,J}(\mathbf{x}_i)$ , possibly with quantifiers, describing  $V_{i,J}$ , for all  $J$  up to some size  $s$ , and we want to get such a formula for  $J$  of size  $s + 1$ .

By the definition, we have the following formula defining  $V_{i,J}$ :

$$F_{i,J}(\mathbf{x}_i) = \left( \bigwedge_{J' \subset J} \tilde{F}_{i,J'}(\mathbf{x}_i) \right) \wedge (\forall \mathbf{x}_j, j \in J) \left[ \left( \bigwedge_{j \in J} \tilde{F}_{j,J \setminus \{j\}}(\mathbf{x}_j) \right) \Rightarrow G_{i,J}(\mathbf{x}_i; \mathbf{x}_j : j \in J) \right],$$

where  $G_{i,J}(\mathbf{x}_i; \mathbf{x}_j, j \in J)$  is the formula asserting that  $\Phi(\mathbf{x}_i; \mathbf{x}_j, j \in J; \mathbf{a}_j, j \in [k] \setminus J \setminus \{i\})$  holds for all  $\mathbf{a}_j \in A_j, j \in [k] \setminus J \setminus \{i\}$ . Now  $G_{i,J}$  can be replaced with a bounded-size formula by the bounded-dimension argument, and this yields the desired bounded-size formula  $\tilde{F}_{i,J}$ , finishing the induction step. We have proved the desired result for algebraic predicates.

Finally, we let  $\Phi = \Phi(\mathbf{x}_1, \dots, \mathbf{x}_n)$  be an arbitrary semialgebraic predicate. We may assume it to be of the form  $\Phi = \Phi_1 \vee \Phi_2 \vee \cdots \vee \Phi_m$ , where each  $\Phi_i$  is a conjunction of polynomial equations and strict inequalities. Given large sets  $A_1, \dots, A_k$  such that  $\Phi$  holds on  $A_1 \times \cdots \times A_k$ , by the  $k$ -partite Ramsey theorem we get that some  $\Phi_\ell$  holds everywhere on  $A'_1 \times \cdots \times A'_k$ , where the  $A'_i \subseteq A_i$  are still large.

Thus, we assume that  $\Phi = \Phi_{=} \wedge \Phi_{<}$ , where  $\Phi_{=}$  is a conjunction of polynomial equations and  $\Phi_{<}$  is a conjunction of strict polynomial inequalities, and that  $\Phi$  holds everywhere on  $A_1 \times \cdots \times A_k$ .

By the argument above for the  $k$ -partite algebraic case, we obtain curves  $Y_1 \supset A_1, \dots, Y_k \supset A_k$  such that  $\Phi_{=}$  holds on  $Y_1 \times \cdots \times Y_k$ . In particular, there are  $\mathbf{a}_i \in A_i, i \in [k]$ , such that each  $\mathbf{a}_i$  is a cluster point of  $Y_i$ . Since  $\Phi_{<}$  holds at  $(\mathbf{a}_1, \dots, \mathbf{a}_k)$ , it also holds for all  $(\mathbf{y}_1, \dots, \mathbf{y}_k)$  from a small neighborhood of  $(\mathbf{a}_1, \dots, \mathbf{a}_k)$ . Thus, we can choose infinite sets  $X_i \subseteq Y_i$  such that  $\Phi$  holds everywhere on  $X_1 \times \cdots \times X_k$  as desired. Theorem 1.9 is proved.  $\square$

As was mentioned in the introduction, the analog of Theorem 1.9 also holds for algebraic predicates over the complex numbers. To see this, only two ingredients in the proof above need to be changed: first, quantifier elimination for the first-order theory of the reals needs to be replaced by quantifier elimination for the first-order theory of algebraically closed fields (see [15] for a recent work and references), and second, we bound the number of connected components of a variety in  $\mathbb{C}^t$  by regarding it as a semialgebraic set in  $\mathbb{R}^{2t}$  and applying the same bounds as in the proof above.

**Algebraic predicates on sequences.** Theorem 1.8 follows easily from the next two lemmas. The first lemma is very similar to Theorem 1.9, but the sets  $A_1, \dots, A_k$  in (i), as well as the sets  $X_1, \dots, X_k$  in (ii), are all equal.

**Lemma 9.1.** *In the setting of Theorem 1.9, if a  $d$ -dimensional  $k$ -ary semialgebraic predicate  $\Phi$ , involving polynomials of degree at most  $D$ , is assumed to hold everywhere on  $A^k = A \times \cdots \times A$ , where  $|A| = N = N(d, k, D)$ , then there is an infinite  $X$  such that  $\Phi$  holds everywhere on  $X^k$ .*

*Proof.* First we assume  $\Phi$  algebraic. We consider the “symmetrization”  $\Phi_{\text{sym}}$  of  $\Phi$ :

$$\Phi_{\text{sym}}(\mathbf{x}_1, \dots, \mathbf{x}_k) := \bigwedge_{\pi \in S_k} \Phi(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(k)})$$

(the conjunction is over all permutations of  $[k]$ ). Then  $\Phi_{\text{sym}}$  also holds everywhere on  $A^k$ . When we define the sets  $V_{i,J}$  for  $\Phi_{\text{sym}}$  as in the proof of Theorem 1.9 above, we have that  $\Phi_{\text{sym}}$ , and hence  $\Phi$ , holds everywhere on  $\prod_{i=1}^k V_{i,[k] \setminus i}$ . By the symmetry of  $\Phi_{\text{sym}}$ , the  $V_{i,J}$  satisfy  $V_{\pi(i), \pi(J)} = V_{i,J}$  for every permutation  $\pi$ , and in particular, the varieties  $V_{i,[k] \setminus i}$  are all equal. Denoting their common value by  $V$ , we get that  $\Phi$  holds everywhere on  $V^k$ . The infinitude of  $V$  follows by bounding the number of components as above.

If  $\Phi$  is semialgebraic, then we again proceed as in the proof of Theorem 1.9, observing that the choice of  $X_1, \dots, X_k$  can also be done symmetrically, i.e., with  $X_1 = \cdots = X_k$ .  $\square$

The second lemma shows that for algebraic predicates, we can pass from “holding everywhere on a long sequence” to “holding everywhere on a large Cartesian power”.

**Lemma 9.2.** *For every  $d, k, D, n$  there exists  $N$  (for which an explicit bound can be given) with the following property. Assuming that a  $d$ -dimensional  $k$ -ary algebraic predicate involving polynomials of degree at most  $D$  holds everywhere on a sequence  $\mathbf{a}$  of length  $N$ , with all terms distinct, then there is an  $n$ -element set  $B \subseteq \mathbf{a}$  such that  $\Phi$  holds everywhere on  $B^k$  (here the inclusion  $B \subseteq \mathbf{a}$  means that  $B$  consists of some of the terms of the sequence  $\mathbf{a}$ ).*

*Proof.* The proof is again based on the bounded-dimension argument. Let us set  $\mathbf{b}^{(0)} := \mathbf{a}$ , and let us assume inductively that  $\mathbf{b}^{(j-1)}$  is a subsequence of  $\mathbf{a}$  of length  $n_{j-1}$  such that  $\Phi(\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_k})$  holds for all choices of  $i_1, \dots, i_{j-1} \in [n_{j-1}]$  and  $1 \leq i_j < i_{j+1} < \cdots < i_k \leq n_{j-1}$ .

For  $i = 1, 2, \dots, n_{j-1}$ , let

$$W_i^{(j)} := \left\{ (\mathbf{x}_1, \dots, \mathbf{x}_j) \in (\mathbb{R}^d)^j : (\forall i_{j+1}, \dots, i_k, i < i_{j+1} < \cdots < i_k) \right. \\ \left. \Phi(\mathbf{x}_1, \dots, \mathbf{x}_j, \mathbf{b}_{i_{j+1}}, \dots, \mathbf{b}_{i_k}) \right\}.$$

The inductive assumption gives  $(\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_j}) \in W_i^{(j)}$  for all  $i_1, \dots, i_{j-1} \in [n_{j-1}]$  and all  $i_j \in [i]$ .

We have  $W_1^{(j)} \subseteq W_2^{(j)} \subseteq \cdots$ , and by the bounded-dimension argument, there are only a bounded number of distinct varieties among the  $W_i^{(j)}$ . Thus, there exist  $i_{\min} < i_{\max}$ , with  $i_{\max} - i_{\min}$  large, such that  $W_{i_{\min}}^{(j)} = W_{i_{\max}}^{(j)}$ .

For all  $i_1, \dots, i_{j-1} \in [n_{j-1}]$ , we have  $(\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_j}) \in W_{i_{\max}}^{(j)} = W_{i_{\min}}^{(j)}$  for all  $i_j \in [i_{\min} + 1, i_{\max}]$ , and by the definition of  $W_{i_{\min}}^{(j)}$ , we obtain that  $\Phi(\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_k})$  holds for all  $i_1, \dots, i_j \in [i_{\min} + 1, i_{\max}]$  and all  $i_{j+1}, \dots, i_k$  with  $i_{\min} < i_{j+1} < \cdots < i_k$ . Hence we can let  $\mathbf{b}^{(j)}$  be the subsequence of  $\mathbf{b}^{(j-1)}$  indexed by  $[i_{\min} + 1, i_{\max}]$ . This finishes the induction step.

The proof is concluded by letting  $B$  be the set of elements of  $\mathbf{b}^{(k)}$ .  $\square$

## Acknowledgments

We would like to thank Uri Andrews for help with logical vocabulary and for kindly scanning Rosenthal’s thesis, and Ehud Hrushovski for a discussion. We are grateful to David Rosenthal for permission to reproduce his thesis.

## References

- [1] I. Bárány, Z. Füredi, and L. Lovász. On the number of halving planes. *Combinatorica*, 10:175–183, 1990.
- [2] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*. Algorithms and Computation in Mathematics 10. Springer, Berlin, 2003.
- [3] B. Bukh, P.-S. Loh, and G. Nivasch. One-sided epsilon-approximants. Manuscript in preparation, 2012.
- [4] B. Bukh, J. Matoušek, and G. Nivasch. Stabbing simplices by points and flats. *Discrete Comput. Geom.*, 43(2):321–338, 2010.
- [5] B. Bukh, J. Matoušek, and G. Nivasch. Lower bounds for weak epsilon-nets and stair-convexity. *Isr. J. Math.*, 182:199–228, 2011.
- [6] M. Eliáš and J. Matoušek. Higher-order Erdős–Szekeres theorems. In *Proc. ACM Sympos. Comput. Geom.*, 2012. Also in arXiv:1111.3824.
- [7] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.
- [8] T. Gerken. On empty convex hexagons in planar point sets. In *J. E. Goodman, J. Pach, R. Pollack (eds.), Twentieth Anniversary Volume: Discrete & Computational Geometry*, pages 1–34. Springer, New York, NY, 2008.
- [9] R.L. Graham, B.L. Rothschild, and J. Spencer. *Ramsey Theory*. J. Wiley & Sons, New York, 1990.
- [10] J. Matoušek. *Lectures on Discrete Geometry*. Springer, New York, 2002.
- [11] J. W. Milnor. On the Betti numbers of real algebraic varieties. *Proc. Amer. Math. Soc.*, 15:275–280, 1964.
- [12] W. Morris and V. Soltan. The Erdős–Szekeres problem on points in convex position—a survey. *Bull. Amer. Math. Soc., New Ser.*, 37(4):437–458, 2000.
- [13] C. M. Nicolás. The empty hexagon theorem. *Discr. Comput. Geom.*, 38(2):389–397, 2007.
- [14] O.A. Oleinik and I.B. Petrovskii. On the topology of of real algebraic surfaces (in Russian). *Izv. Akad. Nauk SSSR*, 13:389–402, 1949.
- [15] S. Puddu and J. Sabia. An effective algorithm for quantifier elimination over algebraically closed fields using straight line programs. *J. Pure Appl. Algebra*, 129(2):173–200, 1998.



- [16] D. A. Rosenthal. The classification of the order indiscernibles of real closed fields and other theories. PhD. thesis, University of Wisconsin–Madison, 1981. Available at [http://www.borisbukh.org/rosenthal\\_thesis.pdf](http://www.borisbukh.org/rosenthal_thesis.pdf).
- [17] M. J. Steele. Variations on the monotone subsequence theme of Erdős and Szekeres. In *D. Aldous et al., editors, Discrete Probability and Algorithms, IMA Volumes in Mathematics and its Applications 72*, pages 111–131. Springer, Berlin etc., 1995.
- [18] A. Tarski. *A decision method for elementary algebra and geometry*. Univ. of California Press, Berkeley, CA, 1951.
- [19] R. Thom. Sur l’homologie des variétés algébriques réelles. In S. S. Cairns, editor, *Differential and Combinatorial Topology*, pages 255–265. Princeton University Press, Princeton, NJ, 1965.
- [20] S. Vrećica and R. Živaljević. The colored Tverberg’s problem and complexes of injective functions. *J. Combin. Theory Ser. A*, 61:309–318, 1992.
- [21] H. E. Warren. Lower bound for approximation by nonlinear manifolds. *Trans. Amer. Math. Soc.*, 133:167–178, 1968.